

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12Q 1/68, C12N 15/10, 15/70		A1	(11) International Publication Number: WO 91/17269 (43) International Publication Date: 14 November 1991 (14.11.91)
(21) International Application Number: PCT/US91/03006 (22) International Filing Date: 2 May 1991 (02.05.91)		(74) Agents: GRANAHAN, Patricia et al.; Hamilton, Brook, Smith & Reynolds, Two Militia Drive, Lexington, MA 02173 (US).	
(30) Priority data: 518,156 3 May 1990 (03.05.90)		US	(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).
(71) Applicant: IG LABORATORIES, INC. [US/US] ; One Mountain Road, Framingham, MA 01701 (US).		Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(72) Inventors: LERNER, Terry, J. ; 122 Adena Road, West Newton, MA 02165 (US). LANDES, Gregory, M. ; 19 Indian Meadow Drive, Northborough, MA 01532 (US). WRIGHT, Gabriela, M., J. S. ; 500-13B Woodcroft Parkway, Durham, NC 27713 (US).			

(54) Title: A METHOD FOR MAPPING A EUKARYOTIC CHROMOSOME

(57) Abstract

A method for mapping discrete DNA sequences complementary to regions of a eukaryotic chromosome, and a method for isolating a gene from a eukaryotic organism of interest. The methods are useful for the efficient identification of genetic markers spaced throughout the human genome. Those genetic markers identified by the mapping method of this invention tend to be associated with the 5' ends of genes. The method, therefore, is useful for the isolation of eukaryotic genes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America

A METHOD FOR MAPPING A EUKARYOTIC CHROMOSOMEBackground of the Invention

The human genome consists of a DNA sequence of some 3 billion base pairs carried on 46 chromosomes. This 5 genetic blueprint provides all of the information necessary for the growth, differentiation and maintenance of the vast array of human cells.

The United States has recently announced, as a national objective, the mapping and sequencing of the 10 human genome. The director of this project, James D. Watson, has recently summarized the importance of this project by stating that the interpretation of the data gained through this work "will not only help us to understand how we function as healthy human beings, but 15 will also explain, at the chemical level, the role of genetic factors in a multitude of diseases, such as cancer, Alzheimer's disease, and schizophrenia, that diminish the individual lives of so many millions of people."

20 The construction of genetic linkage maps using restriction fragment length polymorphisms (RFLPs) was first described by Botstein *et al.* (Am. J. Hum. Genet. 32:314 (1980)). Mapping of a genetic locus by linkage analysis could be performed with high efficiency if 25 polymorphic DNA probes were identified at a spacing of approximately every 10 centimorgans (1cM = 1-2 mbp)

- 2 -

throughout the human genome. A current estimate of the number of evenly spaced markers required to provide markers 20cM apart is 300-700. To obtain such a collection of probes with a spacing of approximately 10 cM 5 or less requires cloning, analyzing and mapping a substantial portion of the genome. This is true of any random mapping method because there is a high probability that newly defined markers will map genetically too close to an existing marker to be useful as an additional 10 marker.

A need exists for an efficient method for identifying genetic markers spaced at intervals of 10cM or less throughout the human genome.

Summary of the Invention

15 The subject invention relates, in one aspect, to a method for ordering a set of discrete DNA sequences complementary to a eukaryotic chromosome for physical and genetic mapping. The method involves providing a set of discrete DNA sequences, each discrete DNA sequence being 20 complementary to a region of the eukaryotic chromosome. The order of the discrete sequences on the chromosome is determined by *in situ* hybridization. Discrete DNA sequences which contain a restriction enzyme recognition sequence containing the dinucleotide CpG and a poly- 25 morphic DNA sequence are then identified.

By determining the order of the discrete DNA sequences (usually genomic clones) on the chromosome prior to further characterization, problems associated with random mapping are eliminated. For example, the 30 number of clones which must be analyzed to define markers

- 3 -

of appropriate spacing is reduced significantly as compared to the number necessary when using the classical random mapping approach. This is true as a result of the initial mapping information provided by in situ hybridization. This allows the usual selection of desirably spaced clones. If selecting anonymous clones (as is done in the classical mapping approach), the analysis of a great many clones is necessary before it would be possible to define a set of markers spanning the chromo-
5 some at a desired spacing.
10

In another aspect, the invention relates to a method for isolating a gene from a eukaryotic organism of interest. A DNA library of genomic DNA clones containing insert DNA from the eukaryotic organism of interest is provided. DNA is purified from individual genomic DNA clones contained within the genomic DNA library and digested with at least one restriction enzyme which recognizes and cleaves a nucleotide sequence which contains the dinucleotide CpG. The products of the
15 restriction enzyme digestion are displayed on a gel and the genomic DNA clones having insert DNA which is recognized and cleaved by the restriction enzyme are identified.
20

The method for isolating a gene from a eukaryotic organism is particularly useful when a gene of interest having a known genetic map position has been identified. In this case, individual genomic DNA clones having insert DNA which is recognized and cleaved by the restriction enzyme which recognizes and cleaves a nucleotide sequence
25 which contains the dinucleotide CpG, are labeled and the
30

-4-

map position of the complementary chromosomal region for each clone is determined physically by *in situ* hybridization. Candidate clones are identified as those having a physical map position near the location of the 5 gene of interest as determined by genetic mapping.

Brief Description of the Figures

Figure 1 is a diagrammatic representation of the vector cHC1.

Figures 2A and 2B are diagrammatic representations 10 showing restriction enzyme cleavage sites for 6 NotI-containing cosmids.

Detailed Description of the Invention

The subject invention is based on the discovery of a simple and convenient method for ordering a set of 15 discrete DNA sequences, each discrete DNA sequences being complementary to a region of a eukaryotic chromosome. The ordering of the discrete DNA sequences is useful for physical and genetic mapping. As used herein, the term "ordering" means to establish the linear relationship of 20 the discrete DNA sequences relative to one another on the chromosome, or a portion of the chromosome.

As an initial step, a set of discrete DNA sequences are provided, each discrete DNA sequence being complementary to a region of the chromosome. In a preferred 25 embodiment, the set of discrete DNA sequences is provided as a chromosome specific genomic DNA library. The chromosome specific genomic DNA library can be obtained from any source, or it can be constructed using known

-5-

techniques. Of particular interest are chromosome specific libraries of human origin, although the methods described herein are applicable to all eukaryotes.

5 Somatic cell hybrids (e.g. hamster/human hybrids) are available which contain a single human chromosome. Such libraries are available for purchase from the American Type Culture Collection (Rockville, Md). Total DNA isolated from such hybrid cells can be fragmented, for example by restriction enzyme digestion, and inserted

10 into an appropriate vector. Genomic clones carrying human inserts (as opposed to inserts of hamster origin) are identified by probing the library with human specific probe DNA (e.g. Alu repeat sequences). Alternatively, the desired human chromosome can be isolated from other

15 human chromosomes by the well known flow sorting technique and used to construct the chromosome specific library.

Preferably, the vector used for construction of the genomic DNA library accomodate inserts of greater than 20 kb. Such vectors include, for example, cosmid vectors, bacteriophage vectors, and yeast artificial chromosomes (YACs). Cosmids are cloning vectors which contain bacteriophage lambda cos signals for in vitro packaging, and allow the cloning of DNA fragments ranging from 30 to 25 50 kb. Bacteriophage vectors (e.g., lambda and P1 bacteriophage vectors) can accommodate from 20- to 100 kb of DNA. A general description of large capacity bacteriophage cloning vectors is provided, Sternbers et al. (Proc. Natl. Acad. Sci. USA 87:103-107 (1990)). A 30 general description of the YAC cloning system is provided, for example, by Burke et al. (Science 236:806-812

(1987)). The YAC cloning vectors can accomodate from 100 to 600 kb of DNA.

The order of the discrete DNA sequences on the human chromosome can be determined, for example, by the *in situ* suppression hybridization method. This technique, which has been described by Lichter *et al.* (Hum. Genet. 80:224-234 (1988)) and is the subject of co-pending application serial no. 07/271,609, permits high resolution physical mapping with fluorescently labeled probe sequences hybridized to interphase or metaphase chromosome preparations. In the Exemplification described below, this method is used to order a set of cosmid clones derived from human chromosome 16.

The discrete DNA sequences are then analyzed for the presence of a restriction enzyme recognition sequence which contains the dinucleotide CpG. This analysis can be conducted by digesting cosmid clones with a rare cutting restriction endonuclease and displaying the products on a gel. Such sequences are known to be rare in the human genome and, therefore, they are frequently referred to as rare cutter sequences (enzymes which recognize such sequences are referred to herein as rare cutting restriction endonucleases). For example, one such sequence, which is recognized by the restriction enzyme Not I, occurs at a frequency of approximately 1/500,000 base pairs. This frequency is convenient for mapping purposes because the theoretical spacing corresponds roughly to resolution limits of the *in situ* suppression hybridization method described above.

This analysis is conducted, for example, by purifying insert DNA from the chromosome specific genomic DNA

library. This purified DNA is then subjected to digestion with a rare cutting restriction endonuclease. Preferably the enzyme's recognition sequence comprises a sequence of at least 6 base pairs. For example, the 5 restriction enzyme Not I, which is discussed in the Exemplification, is particularly useful for this purpose.

The products of this digestion are then displayed on a gel. Electrophoretic methods are well known to those skilled in the art. The preferred gel matrix is agarose; 10 the percentage of agarose can be varied within known ranges to optimize resolution.

Those individual genomic DNA clones having insert DNA which is recognized and cleaved by the rare cutting restriction enzyme are identified by comparing the 15 products of the restriction enzyme digest displayed on a gel with the pattern or expected pattern of an appropriate control sample. An appropriate control sample, for example, is a molecular weight marker or markers having a predetermined electrophoretic mobility. 20 This type of restriction enzyme mapping is a fundamental technique which is well known to those of skill in the art.

In a preferred embodiment, the vector itself contains two recognition sequences for a rare cutting 25 restriction endonuclease. These sites flank the DNA insertion site. Such a vector is shown in Figure 1 and described in the Exemplification. Interpretation is facilitated in this case because if, for example, the insert DNA contains no such site, the digestion product, 30 when displayed on a gel, will include a relatively large band of insert DNA, and a faster migrating band

representing the linear vector DNA. If, on the other hand, the insert does contain such a site, the electrophoretic display will include 3 or more distinct bands.

5 In another screening step, cosmid clones containing a polymorphic DNA sequence are identified. The presence of a polymorphic sequence can be identified, for example, by identifying a restriction fragment length polymorphism (RFLP). One way in which this can be done is by
10 digesting DNA from several unrelated individuals with a variety of different restriction enzymes. This DNA is electrophoretically fractionated, and then transferred to a solid support (e.g. nitrocellulose paper or a nylon filter). This DNA is then screened using labeled cosmid
15 clones. In some individuals, a polymorphism is identified as a restriction fragment having a length differing from the corresponding sequence in another individual. When an RFLP is identified, family members of the individual from which the RFLP was identified are
20 analyzed in a similar manner to determine whether or not the RFLP is inherited according to Mendelian principles. An RFLP which is inherited according to Mendelian principles provides a useful marker for genetic mapping.

25 The clones of this invention, therefore, have two essential characteristics: 1) they span regions of the chromosome containing a recognition sequence for a rare cutting restriction endonuclease, and 2) they contain a DNA sequence polymorphism. A clone which satisfies characteristic 1) can be referred to as a linking clone
30 because it would hybridize to (or link) two adjacent fragments from a total digest of chromosome specific

.9.

human DNA with the restriction enzyme which recognizes the CpG containing sequence. To be maximally useful for mapping purposes, the clones should be spaced from one another along the chromosome at a distance of less than

5 10 cM, and optimally 2-5 cM.

The isolation of clones containing recognition sequences for rare cutting restriction endonucleases offers another advantage. It has been reported that the dinucleotide CpG tends to appear in clusters associated 10 with the 5' ends of eukaryotic genes. These clusters, often referred to as HTF islands (an abbreviation for Hpa I tiny fragment islands) are discussed in two review articles by Bird (Nature 321:209-213 (1986); TIG 3:342-347 (1987)). Clones containing such sequences are, 15 therefore, enriched in DNA sequences corresponding to the 5' ends of genes thereby offering a convenient method for isolating and cloning a eukaryotic gene or genes.

The method involves providing a DNA library of genomic DNA clones containing insert DNA from the 20 eukaryotic organism of interest. As described previously, such a library can be constructed, for example, by isolating DNA and cloning fragments of the DNA into an appropriate vector. When preparing such a library for the purpose of isolating a eukaryotic gene, an important 25 consideration is the size of the DNA insert. Eukaryotic genes can contain multiple introns which do not encode any portion of the protein encoded by the gene, but rather, are excised from the transcribed mRNA prior to translation.

30 For example, the factor VIII gene in the human, which encodes the blood-clotting factor deficient in

-10-

hemophilia A, has been reported to span at least 190 kb (Gitschier *et al.*, Nature 312:326 (1986)). Recent reports indicate that the defective gene responsible for Duchenne's muscular dystrophy may span more than 5 1,000,000 base pairs (Monaco *et al.*, Nature 323:646 (1986)). Therefore, in order to minimize the number of clones which must be screened in order to isolate the gene of interest, the insert DNA is preferably greater than 20 kb in length. Any cloning vector which can 10 accommodate insert DNA of 20 kb or greater is useful for the construction of a genomic library to be screened by the method of this invention.

A preferred vector for the construction of the genomic library is a cosmid vector which can accommodate 15 DNA fragments ranging from 30 to 50 kb. As discussed above, bacteriophage vectors or YAC cloning vectors are also useful for this purpose.

DNA from individual genomic DNA clones contained within the genomic DNA library is purified using known 20 techniques. The purification of cosmid DNA is relatively straightforward and well known in the art. The purification of a yeast artificial chromosome is more technically demanding. One way in which a YAC can be purified is to run a DNA sample containing all yeast 25 chromosomes on a low melting agarose gel by pulsed field agarose gel electrophoresis. This electrophoretic technique enables the resolution of very large DNA molecules. The YAC is then isolated from the other DNA bands in the gel and purified from the gel material using 30 known techniques.

-11-

The purified DNA is then subjected to digestion by a rare cutting restriction enzyme. Many rare cutting restriction endonucleases are known to those skilled in the art. Preferably the enzyme's recognition sequence 5 comprises a sequence of at least 6 bases. Especially preferred is the restriction enzyme NotI.

Once a clone containing a rare cutting restriction endonuclease recognition sequence is isolated, the insert must be further characterized to identify the portion of 10 interest containing the gene. This can be done, for example, by restriction enzyme mapping and DNA sequencing. The sequence of the coding region can then be compared with sequences recorded in gene bank data bases. Using this approach, a genetic locus can be 15 assigned to a gene whose sequence, or a portion thereof, has been determined previously.

This method for isolating a gene is particularly useful when attempting to isolate a gene of interest for which no portion of the nucleotide sequence is known and 20 the identity of the encoded protein is unknown. This is the case, for example, in the study of many human genetic disorders. In the case of such human genetic disorders, a gene is known to be responsible for a disease phenotype, but the identity of the defective protein is 25 unknown. In these situations, unless there are associated gross changes in the chromosomal architecture (e.g. deletion, translocation or inversion) which can be detected by cytogenetic methods, efforts toward localization are limited to studies of genetic linkage in 30 families. This type of analysis typically yields a map resolution only to within several million base pairs. By

providing a genomic library which is specific for the chromosome known to carry the defect, it is possible to reduce the number of clones which theoretically must be screened in order to expect to have a high probability of 5 identifying the clone carrying the gene responsible for the defect.

When the genetic map position of a gene of interest is known, the clones identified which contain recognition sequences for rare cutting enzymes can be labeled and 10 hybridized to human chromosomes in situ as described, for example, by Lichter *et al.* (Hum. Genet. 80:224-234 (1988)). Those clones hybridizing near the location of the gene of interest as determined by genetic mapping represent candidate clones which are analyzed to determine 15 whether they, in fact, encode the gene responsible for the disease phenotype. A variety of strategies can be used to determine whether a candidate clone is, in fact, the gene responsible for the disease phenotype. For example, the clone, or a portion thereof, can be 20 labeled with a reporter group and used to study tissue distribution of complementary mRNA. As discussed in the Exemplification which follows, the gene responsible for autosomal dominant polycystic kidney disease (ADPKD) is known to be manifested in kidney cells. Clones which 25 hybridize to mRNA specifically expressed in kidney cells can be selected for further analysis. For example, such a clone can be used to probe cDNA libraries generated from two sources; individuals having autosomal dominant polycystic kidney disease and individuals not having the 30 disease. Both of the genes isolated from these sources

-13-

can then be sequenced, and the nucleotide change(s) responsible for the disease phenotype can then be determined.

EXEMPLIFICATION

5 The Exemplification which follows sets forth a strategy for the isolation and mapping of an ordered set of discrete DNA sequences for the physical and genetic mapping of individual human chromosomes. The goal of this work is to generate a collection of
10 chromosome-specific cosmid clones that: 1) span the chromosome; 2) contain the recognition sequence of the restriction enzyme Not I; and 3) identify Mendelian RFLPs. By providing both genetic and physical mapping data, this ordered set of discrete DNA sequences will
15 serve to integrate existing physical and genetic chromosome maps.

 The overall efficiency of this strategy depends on:
1) the distribution of Not I sites on a given chromosome;
2) the extent to which the corresponding genomic Not I
20 site is cleavable by that enzyme; and 3) the degree to which the Not I containing clones are genetically polymorphic. To address these questions, a collection of cosmid clones spanning Not I sites on human chromosome 16 has been established. Below is presented the results of
25 initial molecular and genetic characterization of six of these clones.

 The ordered set of discrete DNA sequences derived from this model study of chromosome 16 is useful for the

-14-

study of chromosomal abnormalities. For example, autosomal dominant polycystic kidney disease (ADPKD) is a common genetic disorder with a frequency of 1 per 1,000 in populations of European origin. It is an important 5 cause of chronic renal failure in Europe and the United States, accounting for approximately 10% of all long-term kidney dialysis and transplantation.

It is known that the genetic defect which is responsible for the ADPKD phenotype is linked with both 10 α -globin and phosphoglycolate phosphatase (PGP), thus assigning the locus for the disease to the short arm of chromosome 16 (16p). Physical mapping studies have further refined the localization of this gene to a 600 Kd region of 16p13.3. In the Exemplification which follows 15 a strategy for cloning the ADPKD gene is presented along with results which validate the strategy.

Methods and Materials

Cell Lines and DNAs

Human Epstein-Barr virus (EBV)-transformed lympho- 20 blastoid cell lines were from our laboratory collection. The hypomethylated human lymphoblastoid cell line Til-I was the gift of Dr. Susan Lindsay (Lindsay et al., Hum. Genet. 81:252-256 (1989)). The mouse-human chromosome 16 somatic cell hybrid lines have been described previously 25 (Callen, D.F., Ann. Genet. 29:235-239 (1986); Callen et al., Genomics 2:144-153 (1988); Callen et al., Genomics 4:348-354 (1989)). Table 1 shows the source of the translocation, the original karyotype, and the laboratory name of these hybrid cell lines.

-15-

EBV-transformed lymphoblastoid cell lines were grown in Iscove's Modified Dulbecco's Medium supplemented with 10% horse serum. The hybrid cell lines were grown in F12 medium supplemented with 10% fetal calf serum, 5×10^5 M 5 adenine, and 4 μ g/ml azaserine. Mouse cell line A9 was grown in a similar manner.

DNA was isolated from cultured cell lines by standard phenol/chloroform extraction following proteinase K digestion. High molecular weight DNA for cosmid 10 library construction or pulsed field gel electrophoresis was prepared as described below.

-16-

TABLE 1
HYBRID PANEL FOR CHROMOSOME 16

	<u>Hybrid Line</u>	<u>Portion 16 Present</u>	<u>Other Human Chromosomes Present</u>	<u>Reference</u>
5	CY2	q24->qter	Xpter->Xq26	1
	CY5	q22.15->qter	10pter->10q26	1
	CY6	q22.12->qter	10pter->10q24,+8	2
	CY7*	q13->qter	3pter->3q13.2,+10,+12	1
10	CY8*	q13->qter	11pter->11q14,+4,+7 +8,+20,+21	3
	CY12	p11.2 or p12.2->qter	(12q24>12qter), 1,2,7,8,12,21	6
	CY11	p13.11->qter	11q21->11qter	4
15	CY13	p13.11->qter	1q44->1qter,+3,+11, +14,+17,+20,+21,+22	1
	CY19**	p13.1->qter	13q12.1->13qter	6
	CY14	p13.3->qter	4q31->4qter,+1,+4,+12 +14,+20,+21	5,6
	CY18	intact 16	-	1
20	* The breakpoint of CY7 is distal to the breakpoint of CY8. Both these breakpoints are in 16q13.			
	** The identification of additional human material was not possible because of the presence of unidentifiable human marker chromosomes and translocations between mouse and human chromosomes.			
25				

-17-

TABLE 1 (Continued)

1. Callen, D.F., A mouse/human hybrid cell panel for mapping human chromosome 16. Ann. Genet. 29:235-239 (1986).
- 5 2. Callen, D.F., Hyland, V.J., Baker, E.H., Fratini, A., Simmers, R.N., Mulley, J.C., Sutherland, G.R., Fine mapping of gene probes and anonymous DNA fragments to the long arm of chromosome 16, Genomics 2:144-153 (1988).
- 10 3. Callen, D.F. (unpublished).
4. Koeffler, H.P., Sparkes, R.S., Stang, H., Mohandas, T., Regional assignment of genes for human alpha-globin and phosphoglycolate phosphatase to the short arm of chromosome 16. Proc. Natl. Acad. Sci. 78:7015-7018 (1981).
- 15 5. Derived from cell line of Breuning, M.H., Medan, K., Verjaal, M., Wijnen, J.T., Meera Klou, P., Pearson, P.L., Human globin maps to pter-p13.3 in chromosome 16 distal to PGP, Hum. Genet. 76:287-289 (1987).
- 20 6. Callen, D.F. et al., Mapping the Short Arm of Human Chromosome 16, Genomics 4:348-354 (1989).

Vector and Cosmid Library Construction

The cosmid cloning vector cHCl was derived from the high-copy number, double cos vector c2XBHC (Bates and Swift, Gene 26:137-146 (1983); Bates, P.F., Methods in Enzymology 153:82-94 (1987)) and the "walking easy" vector pWE15 (Stratagene, LaJolla, CA). The small NotI fragment of pWE15, which contains the T3 and T7 promoters and the BamHI cloning site, was enzymatically inserted into a derivative of c2XBHC (gift of Dr. Paul Bates) encoding a single NotI site. The NotI site was created at the single BamHI site of the c2XBHC vector by linker ligation.

-18-

A cosmid library was constructed as described by Swift and Bates (Gene **26**:137-146 (1983)) using cell line CY18 (Callen, D.F., Ann. Genet. **29**:235-239 (1986)) and vector cHC1. High molecular weight DNA was isolated from 5 CY18 cells by proteinase K digestion and very gentle phenol/chloroform extraction followed by dialysis. The resultant DNA was greater than 150 kb as judged by pulsed field gel electrophoresis. The vector was digested with SmaI, dephosphorylated with calf intestinal phosphatase 10 (Boehringer Mannheim), and digested with BamHI. The insert was partially digested with Sau3A and similarly dephosphorylated. Ligation was carried out using 1 μ g of vector arms and 1.5 μ g of target DNA in a final volume of 5 μ l. Reactions were incubated with 200 units of T4 DNA 15 Ligase (New England Biolabs) at room temperature for 4 hours. The DNA was packaged using Gigapack Plus I (Stratagene, LaJolla, CA), titered on the host 1046 (Cami *et al.*, Nucl. Acids. Res. **5**:2381-2390 (1978)), and plated on LB agar plates containing 50 μ g/ml ampicillin.

20 Library Screening

Colonies were plated at low density (1,000 colonies/ 150 mm^2 plate) onto LB agar containing 50 μ g/ml ampicillin. The colonies were transferred onto nylon membrane disks in duplicate and processed, as described by 25 Dillela and Woo (Meth. Enzymol. **152**:199-212 (1987)). Colony filters were probed with nick-translated ³²P-labeled human DNA ($0.5-1 \times 10^6$ cpm/ml). Hybridizations and washes were performed as described for Southern blots (see below). Autoradiography was done with Kodak XAR-5 30 film and an intensifying screen overnight at -70°C. DNA

-19-

was prepared from the human clones using the rapid boiling miniprep procedure (Holmes and Quigley, Anal. Biochem. 114:193-197 (1981)), digested with NotI, and electrophoresed on an 0.7% agarose gel in 1XTBE buffer.

5 For chromosome walking, colony filters were prepared as described above. Radiolabeled RNA probes were transcribed from the bacteriophage T3 and T7 RNA promoters present in the vector (ref) and hybridized at $1-10 \times 10^6$ cpm/ml. Total torula RNA (0.2 mg/ml) was added as
10 competitor.

Restriction Mapping

Cosmid clones were mapped for the rare cutting enzymes BssHII, MluI, NotI, NruI, PvuI and SacII. All enzymes were obtained from New England Biolabs (Beverly, MA) and digestions were performed according to manufacturer's recommendations. Mapping was performed by the single and double enzyme digestion method or partial digestion method of Smith and Birnstiel (Nucl. Acids Res. 3:2387-2398 (1976)), using labeled oligonucleotides 20 to the T₃ and T₇ RNA promoter sequences bordering the insert. Digests were run on two types of gels to optimize sizing: 0.7% agarose in 1XTBE overnight at 60V or 0.4% agarose in 1XTAE at 40V for 16-40 hours. Blotting was done bi-directionally to allow for accurate 25 comparison between hybridizations.

Southern Blot Analysis

Five μ g of genomic DNA were digested with excess restriction enzyme and fractionated on an 0.8% agarose gel in 1XTBE. The DNA in the gel was nicked by partial

-20-

depurination in 0.25 M HCl, denatured in 0.5 M NaOH/1.5 M NaCl, and transferred to nylon membrane (e.g., Magnagraph from MSI, Westboro, MA or SureBlot from Oncor, Gaithersburg, MD). After transfer, the membrane was rinsed in

5 2XSSC, air-dried, and baked in vacuo for 2 hours at 80°C.

³²P-labeled probes were prepared either by nick translation (Rigby et al., J. Mol. Biol. 113:237-251 (1977)) or random priming (Feinberg and Vogelstein, Anal. Biochem. 132:6-13 (1983); Feinberg and Vogelstein, 10 Addendum, 137:266-277 (1984)). Repetitive elements present in some probes were competed out by pre-annealing of the labeled probe with excess sonicated human placental DNA (Scambler et al., Nucl. Acids Res. 15:3639-3651 (1987)). Hybridizations were carried out as

15 described by Church and Gilbert (Proc. Natl. Acad. Sci. USA 81:1991-1995 (1984)). Filters were washed at high stringency (0.1XSSC/0.1% SDS at 65°C) unless otherwise stated in the text. Autoradiography was carried out at -70°C with Kodak XAR-5 film and two intensifying screens.

20 RFLP Analysis

RFLP panels contained digests of DNAs isolated from lymphoblastoid cell lines of six unrelated individuals. The initial enzyme set included BglIII, BclI, EcoRI, HindIII, MspI, PstI, PvuII, SacI, and TagI. When a clone

25 failed to identify an RFLP with these initial nine enzymes, the analysis was extended to include EcoRV, HhaI, RsaI, StuI, and XbaI. Southern blot analysis was carried out as described above. Mendelian inheritance was confirmed in seven 2-generation Caucasian families.

-21-

Pulsed Field Gel Electrophoresis

High molecular weight cellular DNA was encased in agarose blocks as described by Hermann *et al.* (Cell 48:813-825 (1987)). Each 100 μ l block contained the DNA 5 from 2×10^6 IG138 cells or TIL-1 cells, approximately 10 μ g. Restriction digests were carried out as described by Anand, R. (TIG Nov. 278-283 (1986)). Half blocks samples were analyzed by contour clamped homogenous electric field (CHEF) electrophoresis (Chu *et al.*, Science 10 234:1582-1585 (1986)) using a custom-made apparatus (OWL Scientific Plastics, Inc., Cambridge, MA). The gel composition was 0.7% FastLane agarose (FMC Bioproducts, Rockland, ME) in 0.5 x TBE buffer. Electrophoresis was carried out in 0.5XTBE buffer for 16 hours at 180V with a 15 switching interval of 60 seconds. The temperature was maintained at 15°C. Size markers were chromosomes of S. cerevisiae and lambda ladders purchased from Bio-Rad (Burlingame, CA). Southern blot analysis was carried out as described above.

20 In Situ Hybridization

Fluorescent *in situ* hybridization was carried out as detailed by Lichter *et al* (Science 247:64-69 (1990)). Metaphase spreads were prepared from normal cultured lymphocytes (46, XY) by standard procedures of colcemid 25 arrest, hypotonic treatment, and acetic acid-methanol fixation. Cosmid probes were prepared by direct nick translation with biotinylated nucleotides. To facilitate probe penetration and to optimize reannealing, the size of the probe DNA was adjusted empirically to a length of

-22-

150-250 nucleotides by varying the DNase concentration in the nick translation reaction. Slide preparations were routinely counterstained with 200 ng/ml 2-phenylindole-dihydrochloride (DAPI) in 2 XSSC for 5 minutes at room 5 temperature and mounted in 20 mM Tris-HCl (pH 8.0)/90% glycerol containing 2.3% antifade 1,4-diazabicyclo-2(2.2.2) octane. Preparations were visualized on a Zeiss photomicroscope equipped for DAPI and FITC epifluorescence optics, as well as conventional bright field 10 microscopy. Photographs were taken with Kodak Ektachrome 400 (color) film.

RESULTS

Library Construction and Screening

Cosmid vector cHCl shown in Figure 1 is 6 kb in size 15 and has a cloning capacity of 35 to 50 kb. As in its pWE parent, the T3 and T7 promoters flanking the BamHI cloning site allow synthesis of end-specific RNA probes for chromosome walking and mapping. The NotI sites flanking the cloning site allow excision of the cloned 20 DNA insert. Starting with 1.5 μ g of CY18 genomic DNA, we used cHCl to construct a cosmid library of the cell line CY18. CY18 is a mouse-human somatic cell hybrid containing human chromosome 16 as its only human component. The library construction yielded 1×10^6 independent 25 colonies with an average insert size of 41.3 kb.

Approximately 1% (94/10,000) of the clones were identified as human by hybridization to radiolabeled total human DNA. Miniprep DNA was prepared from the

-23-

positive colonies and digested with NotI. NotI cuts out the insert to yield two fragments in those clones without internal NotI sites and 3 or more fragments in those clones with internal NotI sites. Out of 94 human clones, 5 20 had internal NotI sites; 15 had a single site; 2 had two sites; 2 had three sites; and 1 had 4 sites. Restriction analysis verified that all 20 clones are independent isolates.

Regional Localization

10 To verify the chromosomal origin of the clones and to gain initial mapping information, the 20 cosmid clones were biotinylated and used as probes in fluorescent in situ hybridization analysis of human metaphase chromosomes. Hybridization was carried out under conditions 15 that suppress signal from repetitive DNA sequences. Chromosome 16 was identified by hybridization with a chromosome 16-specific alpha satellite DNA clone (Oncor) and by its DAPI-staining pattern. Each clone hybridized exclusively to chromosome 16. The results of this 20 analysis show that the 20 clones are not randomly distributed over the chromosome: 9 map to the long arm and 11 to the short arm. Of the 11 short-arm probes, 7 map to 16p13.3 to 16pter. From this collection of 20 clones, we chose six clones from six distinct chromosomal regions 25 for the studies described below: 16-4N (D16S268), 16-14N (D16S273), 16-30N (D16S271), 16-38N (D16S270), 16-129N (D16S272), and 16-132N (D16S269).

-24-

These six clones were also localized with respect to their position on chromosome 16 by hybridization to the somatic cell hybrid mapping panel described in Table 1. Single copy fragments or fragments containing low levels 5 of repetitive sequence elements were isolated from each clone and used as probe. The results of this analysis are summarized in Table 2. As expected, all six clones hybridized to the parental cell line CY18 and at least one other hybrid cell line. None of the probes hybridized to 10 mouse DNA (cell line A9).

-25-

TABLE 2
SONATIC CELL HYBRIDS

	<u>CY14</u>	<u>CY19</u>	<u>CY11</u>	<u>CY13</u>	<u>CY12</u>	<u>CY8</u>	<u>CY7</u>	<u>CY6</u>	<u>CY5</u>	<u>CY2</u>	<u>CY18</u>	<u>mouse</u>	<u>human</u>
16-14N	+	-	-	-	-	-	-	-	-	-	+	-	+
16-129N	+	+	+	+	-	-	-	-	-	-	+	-	+
16-30N	+	+	+	+	+	-	-	-	-	-	+	-	+
16-38N	+	+	+	+	+	+	-	-	-	-	+	-	+
16-132N	+	+	+	+	+	+	+	+	-	-	+	-	+
16-4N	+	+	+	+	+	+	+	+	+	+	+	-	+

Rare Cutter Restriction Map

Figure 2 shows the restriction enzyme maps for the set of 6 NotI-containing cosmids and overlapping clones isolated by chromosome walking (designated by "W" in the 5 clone name). The maps place the sites for the rare cutting restriction enzymes BssHII, MluI, NotI, NruI, PvuI, and SacII. The clustering of sites for these enzymes is indicative of CpG-rich HTF islands. The NotI sites present in cosmids 16-4N, 16-30N and 16-129N are in 10 close proximity to 2 or more rare cutting restriction enzyme sites, and are most likely island-related. This is not the case for the NotI sites present in the other cosmids. One possible explanation is that these NotI sites are situated in CpG-rich regions which do not 15 encode sites for these or other rare cutting restriction enzymes. HTF island-like regions lacking NotI sites are present in cosmids 30N and 132N.

Linking Clones

To establish the methylation status of each cloned NotI site in genomic DNA and, thus, to identify NotI linking clones, we hybridized the cosmids or cosmid-
5 derived probes to Southern blots containing EcoRI and EcoRI/NotI digests of genomic DNA. The source of DNA for these experiments was either IG138, a human lymphoblast cell line, or Til-1, a hypomethylated cell line recently described by Lindsay *et al.* (Hum. Genet. 81:252-256
10 (1989)). The results from this study are summarized in Table 3.

- 27 -

TABLE 3

<u>Cosmids</u>	<u>Number of</u> <u>NotI Sites</u>	<u>The Extent of Cleavage at Specific</u> <u>NotI Sites in Genomic DNA.</u>	
		<u>138</u>	<u>TIL-1</u>
4N	1	0%	0%
14N	2	0% at site 1 50% at site 2	40% at site 1 80% at site 2
30N	2	100% at sites 1 and 2	100% at sites 1 and 2
30NW6	1	100% at site 2	100% at site 2
38N	1	90%	100%
129N	1	70%	70%
132N	1	0%	0%

The 6 cosmids examined define 6 different loci containing 8 NotI sites. The NotI sites in the loci defined by 16-30N and 16-38N are unmethylated and digest to completion. In contrast, the NotI sites in the loci defined by 16-4N and 16-132N are not cleaved and, therefore, probably methylated at this site in both cell lines. Intermediate extents of methylation are observed

at the NotI sites present in the loci defined by cosmids 16-14N and 16-129N. With the exception of the NotI sites in 16-14N, the extent of methylation of the specific NotI sites is the same for IG138 and TIL-1 DNA. In summary,
5 four of the six clones (16-14N, 16-30N, 16-38N, and 16-129N) can be effectively used to link NotI restriction fragments for long-range physical mapping.

To demonstrate this, we used whole cosmids as probes against Southern blots of NotI-digested genomic DNA
10 fractionated by CHEF gel electrophoresis. With optimal resolution and sufficient single-copy sequence content, this analysis should allow us to identify the NotI fragments on both sides of the NotI site. Figure 3 shows the results of hybridizing the six cosmid clones to CHEF
15 blots of NotI-digested DNA.

Cosmid 16-30N contains 2 NotI sites which lie 25 kb apart and are unmethylated in genomic DNA. We expect 16-30N to anneal to 3 genomic NotI fragments, but in this experiment, we see only 2 NotI fragments (25 kb and 105
20 kb). In the case of 16-38N, whole cosmid hybridizes to a single, resolvable NotI fragment of 150 kb. The location of the NotI site in cosmid 38N should have allowed for the detection of both the leftward and rightward NotI fragments. We conclude that either the leftward fragment
25 ran off the CHEF gel or that the missing NotI fragments is 1600 kb or greater and consequently, unresolved using these electrophoretic conditions. Cosmid 16-129N encodes a single NotI site which is substantially unmethylated (70% cleavage) in the DNA of IG138 and TIL-1 cells.
30 Whole cosmid hybridization to CHEF blots reveals 2

.29.

resolvable NotI fragments of 735 and 200 kb. These fragments may be contiguous or represent partial digestion products.

Cosmid 16-132N encodes a NotI site which is fully 5 methylated in genomic DNA and hybridizes to a fragment which is not resolved using these electrophoretic conditions. Cosmid 16-4N hybridizes to a single NotI fragment consistent with the fully methylated state of the NotI site encoded by this cosmid. The two loci 10 defined by these two cosmids form an interesting comparison since neither of their genomic NotI sites are cleavable, yet the cosmid 16-4N NotI site appears to be in an HTF-island while the NotI site encoded by cosmid 132N exists as an isolated rare cutting restriction 15 enzyme site. Cosmid 16-14N hybridizes to a 160 kb NotI in Till DNA but not to an unresolved fragment in IG138 DNA. This example probably reflects local methylation differences between the two cell lines at this locus.

RFLP Analysis

20 Four of the six clones recognize RFLPs. Mendelian inheritance was demonstrated in 7 2-generation families. This information is summarized in Table 4. 4-15, a 2.4 kb fragment of cosmid 16-4N recognizes SacI and XbaI RFLPs. Although cosmid 16-4N was not used as probe in 25 this study, cosmid 16-4NW1, isolated by chromosome walking from this locus, recognizes EcoRI and BglII RFLPs. All four RFLPs at the 4N locus identify a single haplotype in the 7 families studied. 14-3, a 1.5 kb fragment of cosmid 16-14N, recognizes a PstI polymorphism

-30-

with at least 8 alleles. Because of its location with respect to the PKD-1 gene, this marker may prove useful in testing for autosomal dominant polycystic kidney disease (ADPKD). 129-16, a 5 kb fragment of cosmid 5 16-129N, recognizes a PvuII RFLP. Cosmid 16-132N recognizes an EcoRI RFLP when the whole cosmid is used as probe.

Cosmid clones 16-30N and 16-38N failed to identify polymorphisms with the 14 restriction enzymes used in 10 this study. A cosmid derived from a 34 kb walk from the 16-30N locus (16-30NW6) also failed to identify any RFLPs, as did fragments isolated from these clones. Clone 16-38N hybridizes with the oligonucleotide (CA)₁₀ and, thus, may identify a microsatellite polymorphism (). 15 Clone 16-30N does not hybridize with this oligonucleotide.

- 31 -

TABLE 4

<u>Probe</u>	<u>Locus</u>	<u>Enzyme</u>	<u>Allele (KB)</u>	<u>Frequency</u>
4-15	D16D268	SacI	9	.21
			4.6	.79
5		XbaI	16	.21
			9	.79
16-				
132N	D16S269	EcoRI	17	.54
			14.5	.46
10			2.5	
129-				
16	D16S272	PvuII	8	.57
			6	.43
			2.6	
15 14-3	D16S273	PstI	2.2	.21
			1.3	.035
			1.2	.25
			1.1	.32
			1.05	.07
20			.9	.035
			.84	.035
			.7	.035

Summary

An intermediate goal of this project was to identify Not I linking clones that could serve both as physical and genetic markers for the mapping of a chromosome. Of 5 the 6 cosmid clones selected for characterization, 4 were determined to be Not I linking clones, and of these 4 linking clones, two detect RFLPs. Thus, one-third of the clones yield all of the desired information. To increase the efficiency of the screening process, a linking 10 library can be constructed which contains only Not I sites that are cleavable in genomic DNA. Together, the two libraries should produce a sufficiently large pool of linking clones in which to search for genetic markers.

Clone 16-14N (D16S272) maps to 16p13.3-16p13.13 and 15 thus represents a candidate PKD-1 clone. A variety of strategies can be adopted to determine whether this candidate clone is, in fact, the gene responsible for the disease phenotype. For example, clone 16-14N, or a portion thereof, can be labeled with a reporter group and 20 used to study tissue distribution of complementary mRNA. For example, the gene responsible of ADPKD is manifested in kidney cells. Clones which hybridize to mRNA specifically expressed in kidney cells can be selected for further analysis. For example, such a clone (or a 25 portion thereof) can be used to probe cDNA libraries generated from two sources; individuals having autosomal dominant polycystic kidney disease and individuals not having the disease. Both of the genes isolated from these sources can then be sequenced, and the nucleotide 30 change(s) responsible for the disease phenotype are determined.

- 33 -

The other clones discussed above which do not map to the ADPKD locus can be analyzed by determining their DNA sequence and comparing that sequence with the sequences recorded in gene bank data bases. Using this approach, a 5 genetic locus can be assigned to proteins of known sequence, as well as those of unknown sequence.

- 34 -

CLAIMS

1. A method for ordering a set of discrete DNA sequences for physical and genetic mapping, the method comprising:
 - 5 a) providing a set of discrete DNA sequences, each discrete DNA sequence being complementary to a region of a eukaryotic chromosome;
 - b) determining the order of the discrete sequences on the chromosome by *in situ* hybridization; and
 - 10 c) identifying discrete DNA sequences which contain a restriction enzyme recognition sequence containing the dinucleotide CpG and a polymorphic DNA sequence.
2. A method of Claim 1 wherein the set of discrete DNA sequences complementary to a chromosome is a chromosome specific genomic DNA library.
 - 15 3. A method of Claim 2 wherein the chromosome specific genomic DNA library is a cosmid library.
 4. A method of Claim 2 wherein the chromosome specific genomic DNA library is constructed in a bacteriophage vector.
 - 20 5. A method of Claim 2 wherein the chromosome specific genomic DNA library is constructed in a yeast artificial chromosome.

- 35 -

6. A method of Claim 2 wherein the chromosome specific genomic DNA library is a cosmid library containing inserts from human chromosome 16.
7. A method of Claim 1 wherein the restriction enzyme 5 recognition sequence is recognized by the restriction enzyme Not I.
8. A method of Claim 1 wherein the DNA sequence polymorphism is detectable as an RFLP.
9. A cosmid clone containing:
 - 10 a) a DNA sequence which contains a recognition sequence for a restriction endonuclease which contains the dinucleotide CpG; and
 - b) a DNA sequence polymorphism.

- 36 -

10. A method for isolating a gene from a eukaryotic organism of interest, the method comprising:
 - 5 a) providing a DNA library of genomic DNA clones containing insert DNA from the eukaryotic organism of interest;
 - b) purifying DNA from individual genomic DNA clones contained within the DNA library and digesting the purified DNA with at least one restriction enzyme which recognizes and cleaves 10 a nucleotide sequence which contains the dinucleotide CpG;
 - c) displaying the products of the restriction enzyme digestion reaction on a gel; and
 - d) identifying genomic DNA clones 15 having insert DNA which is recognized and cleaved by the restriction enzyme of step b).
11. A method of Claim 8 wherein the genomic DNA library is constructed within a yeast artificial chromosome.
12. A method of Claim 8 wherein the chromosome specific 20 genomic DNA library is constructed in a bacteriophage vector.
13. A method of Claim 8 wherein the genomic DNA library is constructed in a cosmid vector.
14. A method of Claim 13 wherein the genomic DNA library 25 is constructed within the cosmid vector of Figure 1.

- 37 -

15. A method of Claim 8 wherein the restriction enzyme recognizes and cleaves a DNA sequence of 6 or more base pairs.

16. A method of Claim 8 wherein the restriction enzyme
5 is Not I.

17. A method of Claim 8 wherein the eukaryotic organism of interest is a human.

18. A method of Claim 17 wherein the genomic library is chromosome specific.

19. A method for isolating a gene of interest having a known genetic map position from a eukaryotic organism of interest, the method comprising:
 - a) providing a DNA library of genomic DNA clones containing insert DNA from the eukaryotic organism of interest;
 - b) purifying DNA from individual genomic DNA clones contained within the DNA library and digesting the purified DNA with at least one restriction enzyme which recognizes and cleaves a nucleotide sequence which contains the dinucleotide CpG;
 - c) displaying the products of the step b) on a gel;
 - d) identifying individual genomic DNA clones having insert DNA which is recognized and cleaved by the restriction enzyme of step b);
 - e) labeling clones identified in step d) with a reporter group and determining the map position of the complementary chromosomal region for each clone by *in situ* hybridization; and
 - f) identifying a candidate clone as one which hybridizes near the location of the gene of interest as determined by genetic mapping.
20. A method of Claim 19 wherein the genomic DNA library is constructed within a yeast artificial chromosome.
21. A method of Claim 19 wherein the genomic DNA library is constructed in a cosmid vector.

- 39 -

22. A method of Claim 19 wherein the genomic DNA library is constructed in a bacteriophage vector.
23. A method of Claim 21 wherein the genomic DNA library is constructed within the cosmid vector of Figure 1.
- 5 24. A method of Claim 19 wherein the restriction enzyme recognizes and cleaves a DNA sequence of 6 or more base pairs.
25. A method of Claim 19 wherein the restriction enzyme is Not I.
- 10 26. A method of Claim 19 wherein the eukaryotic organism of interest is a human.
27. A method of Claim 26 wherein the genomic library is chromosome specific.
- 15 28. A method of Claim 19 wherein the eukaryotic gene is responsible for a disease causing genetic disorder which results from the production of a defective protein and the identity of the defective protein is unknown.
- 20 29. A method of Claim 28 wherein the disease is autosomal dominant polycystic kidney disease.

-40-

30. A cosmid vector containing a site for insertion of DNA from a eukaryotic organism of interest, the insertion site being flanked by the nucleotide sequence CGGGCCG.
- 5 31. The cosmid vector of Figure 1.

1/3

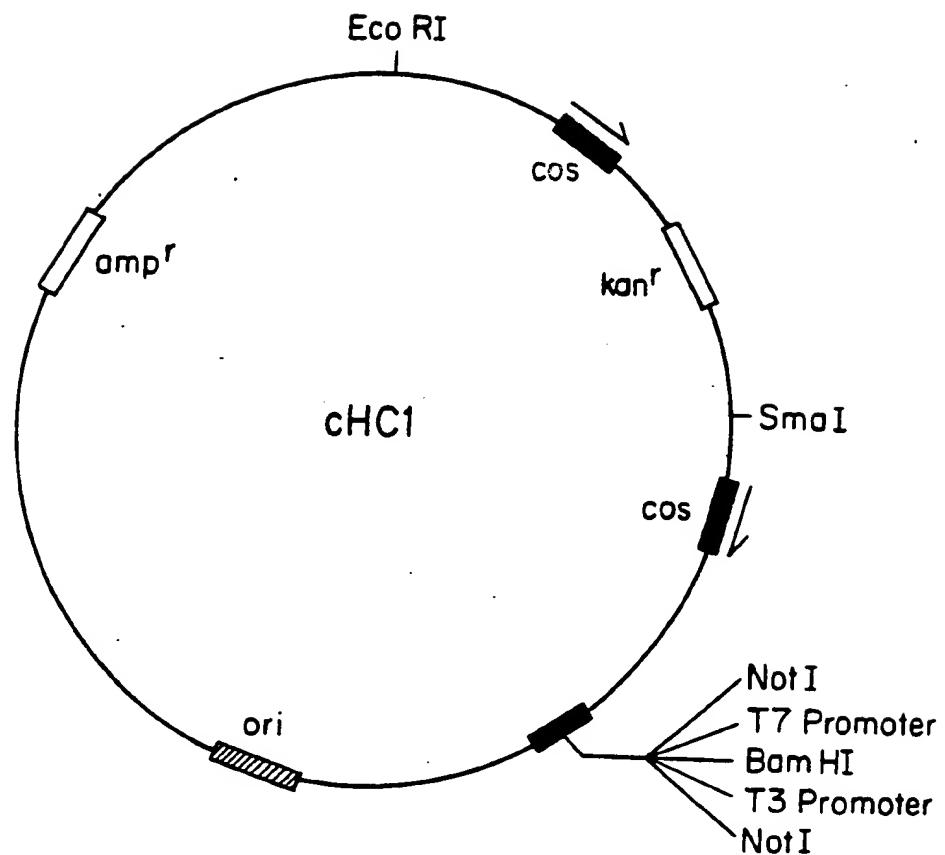
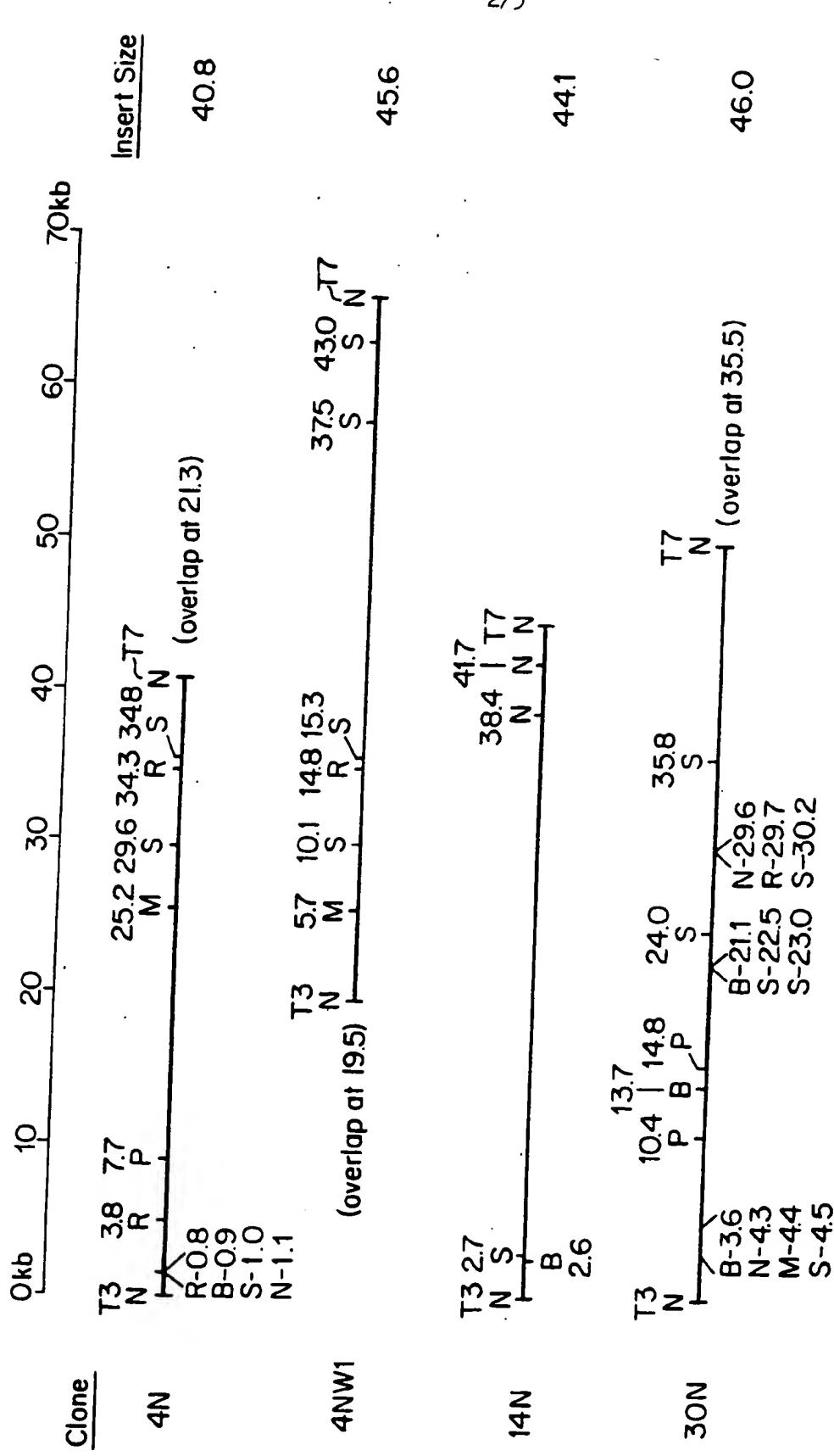
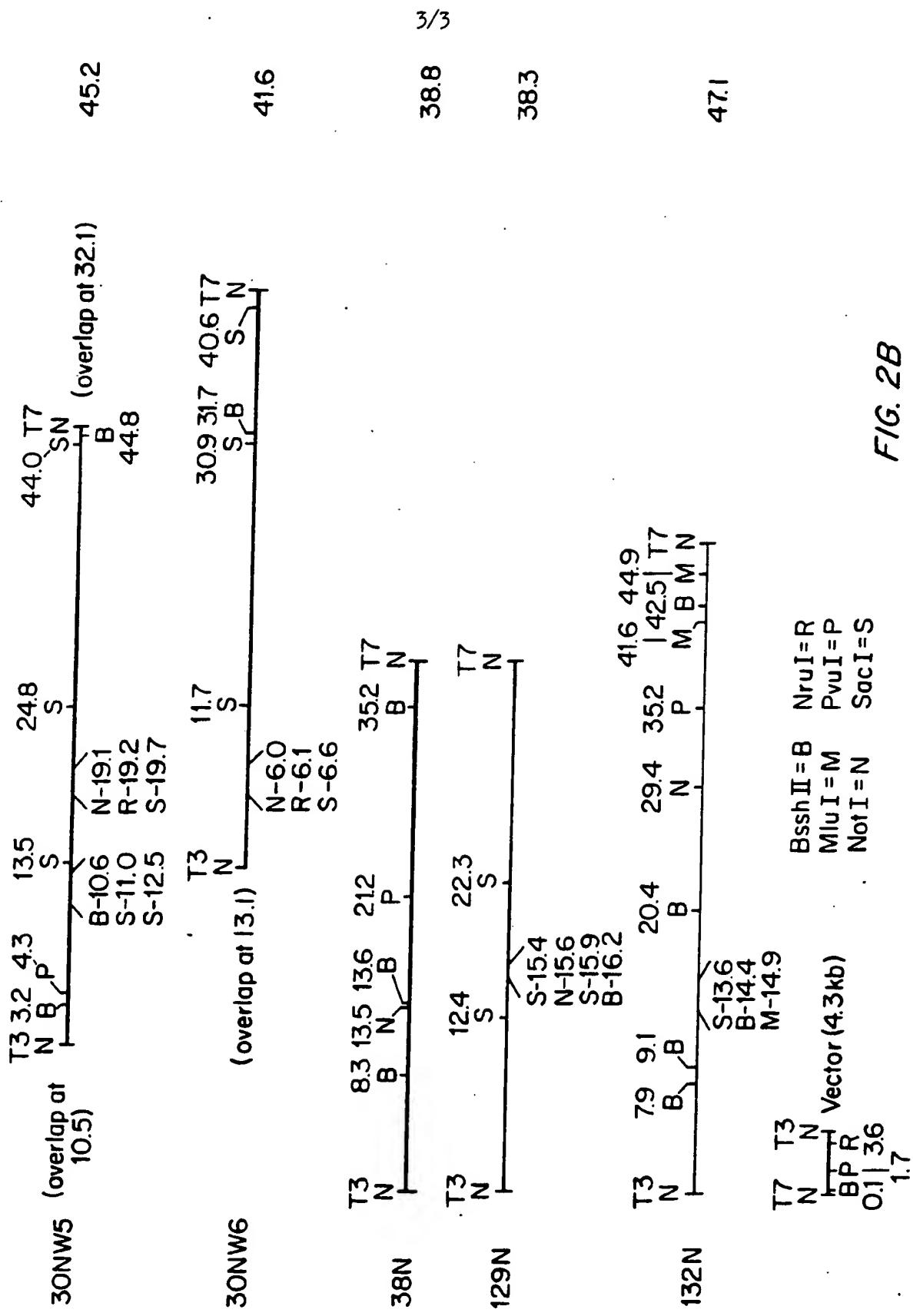


FIG. 1





INTERNATIONAL SEARCH REPORT

International Application No.

PCT/US 91/03006

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all)⁶

According to International Patent Classification (IPC) or to both National Classification and IPC

According to International Patent Classification (IPC), or to World National Classification and IPC
Int. C11,5 C 12 O 1/68 C 12 N 15/10

According to International Patent Classification (IPC), or to be

Classification and IP

C 12 N 15/70

II. FIELDS SEARCHED

Minimum Documentation Searched?

Classification System	Classification Symbols	
Int.C1.5	C 12 Q	C 12 N

Documentation Searched other than Minimum Documentation
to the Extent that such Documents are Included in the Fields Searched⁸

III. DOCUMENTS CONSIDERED TO BE RELEVANT⁹

Category *	Citation of Document, ¹¹ with indication, where appropriate, of the relevant passages ¹²	Relevant to Claim No.13
X	<p>Gene, vol. 79, 1989, Elsevier Science Publishers B.V., (BE), G.A. Evans et al.: "High efficiency vectors for cosmid microcloning and genomic analysis", pages 9-20, see the whole article; especially abstract; page 10, column 1, line 5 - page 10, column 2, line 10; page 13, column 1, line 7 - page 14, line 20</p> <p>---</p>	30,31
Y	<p>---</p>	9
X	<p>---</p> <p>The EMBO Journal, vol. 6, no. 7, 1987, IRL Press Ltd, (Oxford, GB), G.A. Rappold et al.: "Identification of a testis-specific gene from the mouse t-complex next to a CpG-rich island", pages 1975-1980, see the whole article; especially abstract; page 1975, column 2, lines 1-29; page 1977, column 1, line 4 - column 2, line 9; page 1979, column 1, lines 7-39</p> <p>---</p>	10,9

° Special categories of cited documents : ¹⁰

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (see specification).

"O" document referring to an oral disclosure, use, exhibition or other means

P document published prior to the int later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search 06-09-1991	Date of Mailing of this International Search Report 10 OCT 1991
International Searching Authority EUROPEAN PATENT OFFICE	Signature of Authorized Officer Mme N. KUIPER

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
Y	<p>Genomics, vol. 2, 1988, Academic Press, Inc., (San Diego, US), D.F. Callen et al.: "Fine mapping of gene probes and anonymous DNA fragments to the long arm of chromosome 16", pages 144-153, see abstract; page 144, column 1, line 32 - page 145, column 1, line 5; page 145, column 2, lines 3-21; page 150, column 2, line 14 - pages 152, column 2, line 25</p> <p>---</p>	1,6
Y	<p>Nature, vol. 325, 22 January 1987, MacMillan, (London, GB), A. Poustka et al.: "Construction and use of human chromosome jumping libraries from NotI-digested DNA", pages 353-355, see the whole article</p> <p>---</p>	1,6
Y	<p>Science, vol. 247, 5 January 1990, P. Lichter et al.: "High-resolution mapping of human chromosome 11 by in situ hybridization with cosmid clones", pages 64-69, see the whole article; especially abstract; page 66, column 2, line 17 - page 67, column 2, line 4</p> <p>---</p>	1
Y	<p>Human Genetics, vol. 72, no. 1, January 1986, Springer-Verlag, (Berlin, DE), N.E. Buroker et al.: "Four restriction fragment length polymorphisms revealed by probes from a single cosmid map to human chromosome 12q", pages 86-94, see abstract; page 87, column 1, line 5 - page 90, column 1, line 6; page 93, column 1, line 51 - column 2, line 16</p> <p>---</p>	9
A	<p>---</p>	1
A	<p>Human Genetics, vol. 74, no. 4, 1986, Springer-Verlag, (Berlin, DE), L. Bufton et al.: "A highly polymorphic locus on chromosome 16q revealed by a probe from a chromosome-specific cosmid library", pages 425-431, see abstract; page 425, column 1, line 1 - column 2, line 31; page 427, column 1, line 8 - column 2, line 9</p> <p>---</p>	1,6
A	<p>Trends in Genetics, vol. 2, July 1986, Elsevier Science Publishers B.V., (Amsterdam, NL), A. Poustka et al.: "Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics", pages 174-178, see the whole article</p> <p>---</p>	-/-

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)		
Category	Citation of Document, with indication, where appropriate, of the relevant passages	Relevant to Claim No.
P,X	WO,A,9012891 (GENMAP) 1 November 1990, see the whole document; especially page 6, line 9 - page 8, line 21; page 19, line 10 - page 26, line 5; claims 1-48,72 ---	1,30,31
P,Y	Nucleic Acids Research, vol. 18, no. 23, (Oxford, GB), G.A.J. Gillespie et al.: "Cosmid walking and chromosome jumping in the region of PKD1 reveal a locus duplication and three CpG islands", pages 7071-7075, see the whole article; especially abstract -----	1,6

ANNEX TO THE INTERNATIONAL SEARCH REPORT
ON INTERNATIONAL PATENT APPLICATION NO.

US 9103006
SA 48202

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 26/09/91. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A- 9012891	01-11-90	AU-A- 3730689	16-11-90